

#2



OIPE

ENTERED

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/074,596

DATE: 03/01/2002

TIME: 10:54:34

Input Set : A:\Clfr007.app

Output Set: N:\CRF3\03012002\J074596.raw

3 <110> APPLICANT: ROSENBLUM, MICHAEL G.
 4 CHEUNG, LAWRENCE
 6 <120> TITLE OF INVENTION: MODIFIED PROTEINS, DESIGNER TOXINS, AND METHODS OF
 7 MAKING THEEOF
 9 <130> FILE REFERENCE: CLFR:007US
 C--> 11 <140> CURRENT APPLICATION NUMBER: US/10/074,596
 12 <141> CURRENT FILING DATE: 2002-02-12
 14 <150> PRIOR APPLICATION NUMBER: 60/268,402
 15 <151> PRIOR FILING DATE: 2001-02-12
 17 <160> NUMBER OF SEQ ID NOS: 11
 19 <170> SOFTWARE: PatentIn Ver. 2.1
 21 <210> SEQ ID NO: 1
 22 <211> LENGTH: 316
 23 <212> TYPE: PRT
 24 <213> ORGANISM: Gelonium multiflorum
 26 <400> SEQUENCE: 1
 27 Met Lys Gly Asn Met Lys Val Tyr Trp Ile Lys Ile Ala Val Ala Thr
 28 1 5 10 15
 30 Trp Phe Cys Cys Thr Thr Ile Val Leu Gly Ser Thr Ala Arg Ile Phe
 31 20 25 30
 33 Ser Leu Pro Thr Asn Asp Glu Glu Thr Ser Lys Thr Leu Gly Leu
 34 35 40 45
 36 Asp Thr Val Ser Phe Ser Thr Lys Gly Ala Thr Tyr Ile Thr Tyr Val
 37 50 55 60
 39 Asn Phe Leu Asn Glu Leu Arg Val Lys Leu Lys Pro Glu Gly Asn Ser
 40 65 70 75 80
 42 His Gly Ile Pro Leu Arg Lys Lys Cys Asp Asp Pro Gly Lys Cys
 43 85 90 95
 45 Phe Val Leu Val Ala Leu Ser Asn Asp Asn Gly Gln Leu Ala Glu Ile
 46 100 105 110
 48 Ala Ile Asp Val Thr Ser Val Tyr Val Val Gly Tyr Gln Val Arg Asn
 49 115 120 125
 51 Arg Ser Tyr Phe Phe Lys Asp Ala Pro Asp Ala Ala Tyr Glu Gly Leu
 52 130 135 140
 54 Phe Lys Asn Thr Ile Lys Thr Arg Leu His Phe Gly Gly Ser Tyr Pro
 55 145 150 155 160
 57 Ser Leu Glu Gly Glu Lys Ala Tyr Arg Glu Thr Thr Asp Leu Gly Ile
 58 165 170 175
 60 Glu Pro Leu Arg Ile Gly Ile Lys Lys Leu Asp Glu Asn Ala Ile Asp
 61 180 185 190
 63 Asn Tyr Lys Pro Thr Glu Ile Ala Ser Ser Leu Leu Val Val Ile Gln
 64 195 200 205
 66 Met Val Ser Glu Ala Ala Arg Phe Thr Phe Ile Glu Asn Gln Ile Arg

RAW SEQUENCE LISTING
PATENT APPLICATION: US/10/074,596

DATE: 03/01/2002
TIME: 10:54:34

Input Set : A:\Clfr007.app
Output Set: N:\CRF3\03012002\J074596.raw

```

67      210      215      220
69 Asn Asn Phe Gln Gln Arg Ile Arg Pro Ala Asn Asn Thr Ile Ser Leu
70 225      230      235      240
72 Glu Asn Lys Trp Gly Lys Leu Ser Phe Gln Ile Arg Thr Ser Gly Ala
73      245      250      255
75 Asn Gly Met Phe Ser Glu Ala Val Glu Leu Glu Arg Ala Asn Gly Lys
76      260      265      270
78 Lys Tyr Tyr Val Thr Ala Val Asp Gln Val Lys Pro Lys Ile Ala Leu
79      275      280      285
81 Leu Lys Phe Val Asp Lys Asp Pro Lys Thr Ser Leu Ala Ala Glu Leu
82      290      295      300
84 Ile Ile Gln Asn Tyr Glu Ser Leu Val Gly Phe Asp
85 305      310      315
88 <210> SEQ ID NO: 2
89 <211> LENGTH: 1176
90 <212> TYPE: DNA
91 <213> ORGANISM: Gelonium multiflorum
93 <400> SEQUENCE: 2
94 cagcttctca cttgtttggg ataatgaaag ggaacatgaa ggtgtactgg attaagattg 60
95 ctgtggcgac atggttttgc tgcactacta ttgtacttgg atcaacggcg aggattttct 120
96 ctcttccac aaatgatgaa gaagaaacca gtaagacgct tggcctggac accgtgagct 180
97 ttagcactaa aggtgccact tatattacct acgtgaattt cttgaatgag ctacgagtta 240
98 aattgaaacc cgaaggtaac agccatggaa tccattgct gcgcaaaaaa tgtgatgatc 300
99 ctggaaagtg ttctgttttg gtagcgcttt caaatgacaa tggacagttg gcggaatatg 360
100 ctatagatgt tacaagtgtt tatgtggtgg gctatcaagt aagaacacaga tcttacttct 420
101 ttaaagatgc tccagatgct gcttacgaag gcctcttcaa aaacacaatt aaaacaagac 480
102 ttcatthttg cggcagctat ccctcgctgg aaggtagagaa ggcatataga gagacaacag 540
103 acttgggcat tgaaccatta aggattggca tcaagaaact tgatgaaaat gcgatagaca 600
104 attataaacc aacggagata gctagttctc tattggttgt tattcaaag gtgtctgaag 660
105 cagctcgatt cacctttatt gagaaccaa ttagaaataa ctttcaacag agaattcgcc 720
106 cggcgaataa tacaatcagc cttgagaata aatggggtaa actctcgttc cagatccgga 780
107 catcaggtgc aaatggaatg ttttcggagg cagttgaatt ggaacgtgca aatggcaaaa 840
108 aatactatgt caccgcagtt gatcaagtaa aacccaaat agcactcttg aagttcgctg 900
109 ataaagatcc taaaacgagc cttgctgctg aattgataat ccagaactat gagtcattag 960
110 tgggctttga ttagtacaac ttattgtgct ttttatatat tatagatatg atgccgggcc 1020
111 atgtattggc cttcgtagct taaataaagg catcgaatat tagcctcggt ggtgtatcta 1080
112 tcatgctgtg ttgtaaaact gccaatgttt atgttatcaa acagaaattg gcatgaagtt 1140
113 tctgtacaag tgttcaataa actgggctat acatgc 1176
116 <210> SEQ ID NO: 3
117 <211> LENGTH: 33
118 <212> TYPE: DNA
119 <213> ORGANISM: Homo sapiens
121 <400> SEQUENCE: 3
122 gctgcccaac cagccatggc ggacattgtg atg
125 <210> SEQ ID NO: 4
126 <211> LENGTH: 50
127 <212> TYPE: DNA
128 <213> ORGANISM: Homo sapiens
130 <400> SEQUENCE: 4

```

33

RAW SEQUENCE LISTING
PATENT APPLICATION: US/10/074,596

DATE: 03/01/2002
TIME: 10:54:34

Input Set : A:\Clfr007.app
Output Set: N:\CRF3\03012002\J074596.raw

```

131 gccggagcct ggcttgcaacg ctgccgctgg tggagccttt gatcatccag          50
134 <210> SEQ ID NO: 5
135 <211> LENGTH: 45
136 <212> TYPE: DNA
137 <213> ORGANISM: Homo sapiens
139 <400> SEQUENCE: 5
140 aagccaggct cgggcgaagg cagcaccaaa ggcgaagtga aggtt          45
143 <210> SEQ ID NO: 6
144 <211> LENGTH: 30
145 <212> TYPE: DNA
146 <213> ORGANISM: Homo sapiens
148 <400> SEQUENCE: 6
149 gccaccgcca ccactagttg aggagactgt          30
152 <210> SEQ ID NO: 7
153 <211> LENGTH: 51
154 <212> TYPE: DNA
155 <213> ORGANISM: Artificial Sequence
157 <220> FEATURE:
158 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
159     Primer
161 <400> SEQUENCE: 7
162 ggcggtggct ccgtcatgac ggacattgtg atgacccagt ctcaaaaatt c          51
165 <210> SEQ ID NO: 8
166 <211> LENGTH: 33
167 <212> TYPE: DNA
168 <213> ORGANISM: Artificial Sequence
170 <220> FEATURE:
171 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
172     Primer
174 <400> SEQUENCE: 8
175 ggtggcgggtg gtcctgtct agacaccgtg acg          33
178 <210> SEQ ID NO: 9
179 <211> LENGTH: 45
180 <212> TYPE: DNA
181 <213> ORGANISM: Artificial Sequence
183 <220> FEATURE:
184 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
185     Primer
187 <400> SEQUENCE: 9
188 aaggctcgtg tcgacctga gtcattaagc tttaggatct ttatc          45
191 <210> SEQ ID NO: 10
192 <211> LENGTH: 1527
193 <212> TYPE: DNA
194 <213> ORGANISM: Artificial Sequence
196 <220> FEATURE:
197 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
199 <220> FEATURE:
200 <221> NAME/KEY: CDS
201 <222> LOCATION: (1)..(1521)

```

RAW SEQUENCE LISTING
PATENT APPLICATION: US/10/074,596

DATE: 03/01/2002
TIME: 10:54:34

Input Set : A:\Clfr007.app
Output Set: N:\CRF3\03012002\J074596.raw

```

203 <400> SEQUENCE: 10
204 atg acg gac att gtg atg acc cag tct caa aaa ttc atg tcc aca tca 48
205 Met Thr Asp Ile Val Met Thr Gln Ser Gln Lys Phe Met Ser Thr Ser
206 1 5 10 15
208 gta gga gac agg gtc agc gtc acc tgc aag gcc agt cag aat gtg gat 96
209 Val Gly Asp Arg Val Ser Val Thr Cys Lys Ala Ser Gln Asn Val Asp
210 20 25 30
212 act aat gta gcc tgg tat caa caa aaa cca ggg caa tct cct gaa cca 144
213 Thr Asn Val Ala Trp Tyr Gln Gln Lys Pro Gly Gln Ser Pro Glu Pro
214 35 40 45
216 ctg ctt ttc tcg gca tcc tac cgt tac act gga gtc cct gat cgc ttc 192
217 Leu Leu Phe Ser Ala Ser Tyr Arg Tyr Thr Gly Val Pro Asp Arg Phe
218 50 55 60
220 aca ggc agt gga tct ggg aca gat ttc act ctc acc atc agc aat gtg 240
221 Thr Gly Ser Gly Ser Gly Thr Asp Phe Thr Leu Thr Ile Ser Asn Val
222 65 70 75 80
224 cag tct gaa gac ttg gca gag tat ttc tgt cag caa tat aac agc tat 288
225 Gln Ser Glu Asp Leu Ala Glu Tyr Phe Cys Gln Gln Tyr Asn Ser Tyr
226 85 90 95
228 cct ctg acg ttc ggt gga ggc acc aag ctg gag atc aaa ggc tcc acc 336
229 Pro Leu Thr Phe Gly Gly Gly Thr Lys Leu Glu Ile Lys Gly Ser Thr
230 100 105 110
232 agc ggc agc ggt aag cca ggc tcc ggc gaa ggc agc acc aaa ggc gaa 384
233 Ser Gly Ser Gly Lys Pro Gly Ser Gly Glu Gly Ser Thr Lys Gly Glu
234 115 120 125
236 gtg aag gtt gag gag tct gga gga ggc ttg gtg caa cct gga gga tcc 432
237 Val Lys Val Glu Glu Ser Gly Gly Gly Leu Val Gln Pro Gly Gly Ser
238 130 135 140
240 atg aaa ctc tcc tgt gtt gtc tct gga ttc act ttc ggt aat tac tgg 480
241 Met Lys Leu Ser Cys Val Val Ser Gly Phe Thr Phe Gly Asn Tyr Trp
242 145 150 155 160
244 atg aac tgg gtc cgc cag tct cca gag aag ggg ctt gag tgg att gca 528
245 Met Asn Trp Val Arg Gln Ser Pro Glu Lys Gly Leu Glu Trp Ile Ala
246 165 170 175
248 gaa att aga ttg aaa tcc aat aat ttt gca aga tat tat gcg gag tct 576
249 Glu Ile Arg Leu Lys Ser Asn Asn Phe Ala Arg Tyr Tyr Ala Glu Ser
250 180 185 190
252 gtg aaa ggg agg ttc acc atc tca aga gat gat tcc aaa agt agt gtc 624
253 Val Lys Gly Arg Phe Thr Ile Ser Arg Asp Asp Ser Lys Ser Ser Val
254 195 200 205
256 tac ctg caa atg atc aac cta aga gct gaa gat act ggc att tat tac 672
257 Tyr Leu Gln Met Ile Asn Leu Arg Ala Glu Asp Thr Gly Ile Tyr Tyr
258 210 215 220
260 tgt acc agt tat ggt aac tac gtt ggg cac tat ttt gac cac tgg ggc 720
261 Cys Thr Ser Tyr Gly Asn Tyr Val Gly His Tyr Phe Asp His Trp Gly
262 225 230 235 240
264 caa ggc acc act ctc acc gtc tcc tca gct agc ggt ggc ggt ggc tcc 768
265 Gln Gly Thr Thr Leu Thr Val Ser Ser Ala Ser Gly Gly Gly Gly Ser
266 245 250 255

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/074,596

DATE: 03/01/2002

TIME: 10:54:34

Input Set : A:\Clfr007.app

Output Set: N:\CRF3\03012002\J074596.raw

```

268 ggt cta gac acc gtg agc ttt agc act aaa ggt gcc act tat att acc 816
269 Gly Leu Asp Thr Val Ser Phe Ser Thr Lys Gly Ala Thr Tyr Ile Thr
270 260 265 270
272 tac gtg aat ttc ttg aat gag cta cga gtt aaa ttg aaa ccc gaa ggt 864
273 Tyr Val Asn Phe Leu Asn Glu Leu Arg Val Lys Leu Lys Pro Glu Gly
274 275 280 285
276 aac agc cat gga atc cca ttg ctg cgc aaa aaa tgt gat gat cct gga 912
277 Asn Ser His Gly Ile Pro Leu Leu Arg Lys Lys Cys Asp Asp Pro Gly
278 290 295 300
280 aag tgt ttc gtt ttg gta gcg ctt tca aat gac aat gga cag ttg gcg 960
281 Lys Cys Phe Val Leu Val Ala Leu Ser Asn Asp Asn Gly Gln Leu Ala
282 305 310 315 320
284 gaa ata gct ata gat gtt aca agt gtt tat gtg gtg ggc tat caa gta 1008
285 Glu Ile Ala Ile Asp Val Thr Ser Val Tyr Val Val Gly Tyr Gln Val
286 325 330 335
288 aga aac aga tct tac ttc ttt aaa gat gct cca gat gct gct tac gaa 1056
289 Arg Asn Arg Ser Tyr Phe Phe Lys Asp Ala Pro Asp Ala Tyr Glu
290 340 345 350
292 ggc ctc ttc aaa aac aca att aaa aca aga ctt cat ttt ggc ggc agc 1104
293 Gly Leu Phe Lys Asn Thr Ile Lys Thr Arg Leu His Phe Gly Gly Ser
294 355 360 365
296 tat ccc tcg ctg gaa ggt gag aag gca tat aga gag aca aca gac ttg 1152
297 Tyr Pro Ser Leu Glu Gly Glu Lys Ala Tyr Arg Glu Thr Thr Asp Leu
298 370 375 380
300 ggc att gaa cca tta agg att ggc atc aag aaa ctt gat gaa aat gcg 1200
301 Gly Ile Glu Pro Leu Arg Ile Gly Ile Lys Lys Leu Asp Glu Asn Ala
302 385 390 395 400
304 ata gac aat tat aaa cca acg gag ata gct agt tct cta ttg gtt gtt 1248
305 Ile Asp Asn Tyr Lys Pro Thr Glu Ile Ala Ser Ser Leu Leu Val Val
306 405 410 415
308 att caa atg gtg tct gaa gca gct cga ttc acc ttt att gag aac caa 1296
309 Ile Gln Met Val Ser Glu Ala Ala Arg Phe Thr Phe Ile Glu Asn Gln
310 420 425 430
312 att aga aat aac ttt caa cag aga att cgc ccg gcg aat aat aca atc 1344
313 Ile Arg Asn Asn Phe Gln Gln Arg Ile Arg Pro Ala Asn Asn Thr Ile
314 435 440 445
316 agc ctt gag aat aaa tgg ggt aaa ctc tcg ttc cag atc cgg aca tca 1392
317 Ser Leu Glu Asn Lys Trp Gly Lys Leu Ser Phe Gln Ile Arg Thr Ser
318 450 455 460
320 ggt gca aat gga atg ttt tcg gag gca gtt gaa ttg gaa cgt gca aat 1440
321 Gly Ala Asn Gly Met Phe Ser Glu Ala Val Glu Leu Glu Arg Ala Asn
322 465 470 475 480
324 ggc aaa aaa tac tat gtc acc gca gtt gat caa gta aaa ccc aaa ata 1488
325 Gly Lys Lys Tyr Tyr Val Thr Ala Val Asp Gln Val Lys Pro Lys Ile
326 485 490 495
328 gca ctc ttg aag ttc gtc gat aaa gat cct aaa taatga 1527
329 Ala Leu Leu Lys Phe Val Asp Lys Asp Pro Lys
330 500 505
333 <210> SEQ ID NO: 11

```

VERIFICATION SUMMARY

PATENT APPLICATION: US/10/074,596

DATE: 03/01/2002

TIME: 10:54:35

Input Set : A:\Clfr007.app

Output Set: N:\CRF3\03012002\J074596.raw

L:11 M:270 C: Current Application Number differs, Replaced Current Application Number
L:337 M:258 W: Mandatory Feature missing, <220> FEATURE: